BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 8: e134757
doi: 10.3897/biss.8.134757

OPEN ACCESS

Conference Abstract

Using Webby FDOs to Integrate AI Taxon Identification and Citizen Science

Jonas Grieb‡, Claus Weiland‡, Alexander Wolodkin‡, Leyla Jael Castro§, Stian Soiland-Reyes|, Maya Beukes‡, Martin Jansen‡

‡ Senckenberg – Leibniz Institution for Biodiversity and Earth System Research, Frankfurt am Main, Germany
§ ZB MED Information Centre for Life Sciences, Cologne, Germany
| The University of Manchester, Manchester, United Kingdom

Corresponding author: Jonas Grieb (jonas.grieb@senckenberg.de), Claus Weiland (cweiland@senckenberg.de)

Received: 14 Aug 2024 | Published: 15 Aug 2024

Citation: Grieb J, Weiland C, Wolodkin A, Castro LJ, Soiland-Reyes S, Beukes M, Jansen M (2024) Using Webby FDOs to Integrate AI Taxon Identification and Citizen Science. Biodiversity Information Science and Standards 8: e134757. https://doi.org/10.3897/biss.8.134757

Abstract

Camera traps and passive acoustic devices are particularly useful in providing non-invasive methods to document wildlife diversity, ecology, behavior, and conservation. The application of autonomous Internet of Things (IoT) sensors is constantly developing and opens up new application possibilities for research and nature conservation such as taxon identification based on real-time audio processing in the field (Höchst et al. 2022). Furthermore, the amount of associated recorded digital photos, videos, and audio files is growing at a rapid pace, producing too much data to make human annotation feasible.

Machine-learning, in contrast, can generate (baseline) annotations at scale on high-throughput data but may not capture all details compared with the complex contextual understanding of human annotators.

We developed the WildLIVE platform*[1] to effectively combine machine-learning (taxon and individual identification of animals) and “Human-in-the-Loop” citizen science (Fig. 1, Jansen et al. 2024). Our primary objective is to combine crowdsourced curation of digital image, audio, and video content from biodiversity monitoring with (semi-)autonomous data processing by machines, and subsequently mobilize these data as machine-actionable knowledge units, i.e., FAIR Digital Objects (FDO, Wittenburg et al. 2023)

© Grieb J et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

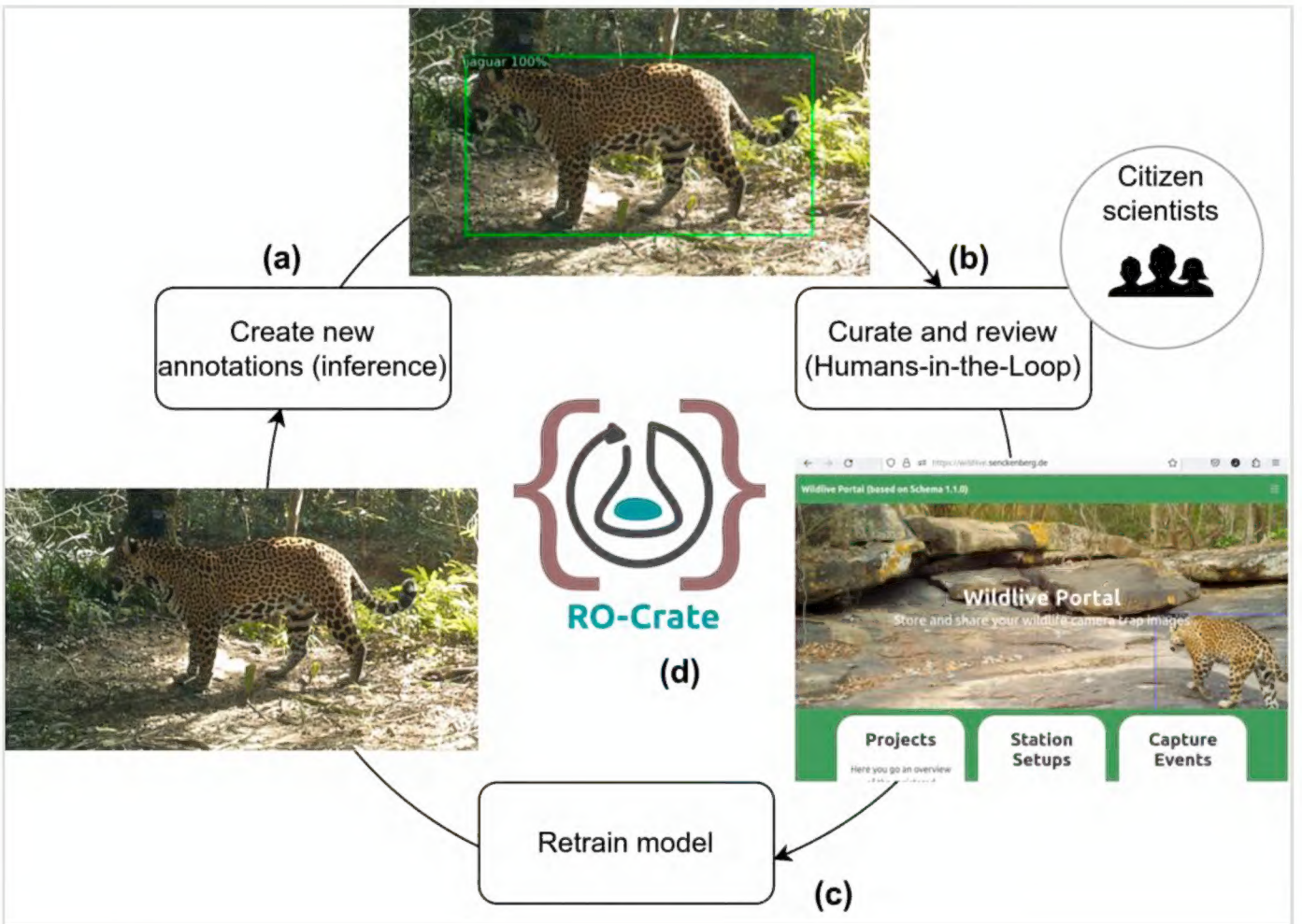

Figure 1.

Data flow in the WildLIVE portal: A machine-learning service provides automatic high throughput annotations as baseline (a). Annotations are then subject to review and refinement by citizen scientists (b). Reviewed data are subsequently compiled into new training data (c). The data model (d), based on lightweight packaging with RO-Crate, captures contextual (provenance) information from all human and machine-based operations.

To this effect, the data model of WildLIVE features a “webby" FDO approach leveraging web-based components involving Research Object Crate (RO-Crate) and FAIR Signposting to enable packaging of an observing process’ contextual information (e.g., metadata of sensors, geolocation, and links to content stream), together with operational semantics giving machines the information needed to autonomously process (e.g., detect regions of interest in images, Younis et al. 2020) the actual data (Soiland-Reyes et al. 2024). To represent the semantics of data capture events, we designed an ontology, the WildLife Monitoring Ontology (WLMO), which provides a formal description of fundamental concepts and relations.

The talk will provide an overview of the platform’s development status and the technology stack employed (combining RO-Crate and FAIR Signposting with AI plus “Humans-in-the-Loop”) for data exchange with emerging data infrastructures such as the Common European Data Spaces.

Keywords

Human-in-the-Loop, deep learning, FAIR Digital Object, RO-Crate, machine-actionability, Wildlife Monitoring Ontology, Common European Data Spaces, camera trap

Presenting author

Claus Weiland

Presented at

SPNHC-TDWG 2024

Funding program

CAMTRAPPER– CAMera TRap Archive and Public Portal for Exploration & Research, Deutsche Forschungsgemeinschaft (DFG) no. 437771903

AI4WildLIVE - Citizen Science Portal, Archive and Analysis Tools for Multimodal Monitoring Data, funding line: BiodivKI - Methods of Artificial Intelligence as an Instrument of Biodiversity Research, German Federal Ministry of Education and Research (BMBF)

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Höchst J, Bellafkir H, Lampe P, Vogelbacher M, Mühling M, Schneider D, Lindner K, Rösner S, Schabo D, Farwig N, Freisleben B (2022) Bird@Edge: Bird Species Recognition at the Edge. Lecture Notes in Computer Science69-86. https://doi.org/10.1007/978-3-031-17436-0_6
- Jansen M, Beukes M, Weiland C, Blumer M, Rudolfi M, Poerting J, Meißner R, Weiß M, Condori Y, Aramayo-Ledezma G, Grieb J, Wolodkin A, Aramayo-Bejarano JL (2024) Engaging Citizen Scientists in Biodiversity Monitoring: Insights from the WildLIVE! Project. Citizen Science: Theory and Practice 9 (1). https://doi.org/10.5334/cstp.665
- Soiland-Reyes S, Castro LJ, Ravinder R, Weiland C, Grieb J, Rogers A, Blanchi C, Van de Sompel H (2024) BioHackEU23 report: Enabling FAIR Digital Objects with RO-Crate, Signposting and Bioschemas. Biohackathon series: BioHackEU23 https://doi.org/10.37044/osf.io/gmk2h
- Wittenburg P, Schwardmann U, Blanchi C, Weiland C (2023) FDOs to Enable Cross-Silo Work. Proceedings of the Conference on Research Data Infrastructure 1 https://doi.org/10.52825/cordi.v1i.263

- Younis S, Schmidt M, Weiland C, Dressler S, Seeger B, Hickler T (2020) Detection and annotation of plant organs from digitised herbarium scans using deep learning. Biodiversity Data Journal 8 https://doi.org/10.3897/bdj.8.e57090

Endnotes

*1 https://wildlive.senckenberg.de